AI Simplified # Chapter 2

#Artificial Narrow intelligence

RECAP : Artificial Narrow intelligence

Mathematics

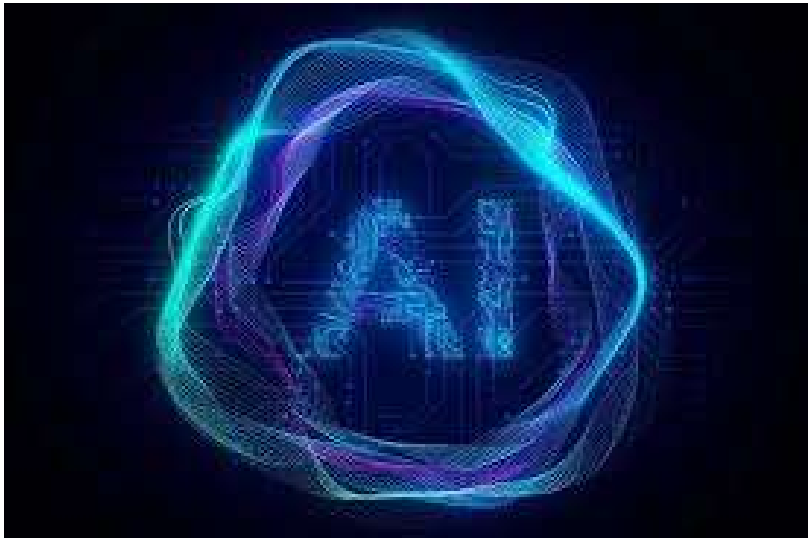
- Vector and Vector spaces
- Vector Operations
- Matrix and matrix operations
- Single Value decomposition
- Probability and Statistics
- Random variable and distribution
- Concept of Hypothesis
- Information theory

Data

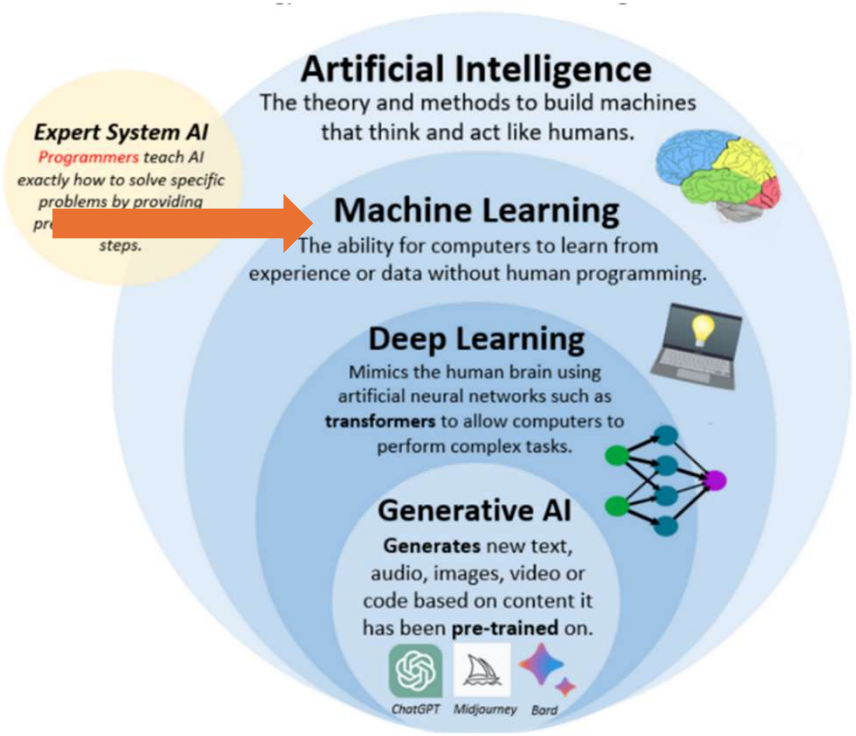
- Structure Data
- Semi structure data
- Text /Image/Video/Audio
- Data formats

Last Week in AI news

- Explainable AI systems for ship navigation have been introduced, which raise trust and decrease human error in maritime operations
- AI tools have been developed to better assess Parkinson's disease and other movement disorders, improving diagnostic accuracy and patient care
- Engineers have advanced toward creating a fault-tolerant quantum computer, which could have profound implications for AI and computing power
- Elon Musk announced the upcoming release of Grok 3, a new AI chatbot claimed to outperform existing models like OpenAI's ChatGPT, highlighting ongoing competition and innovation in conversational AI
- "The Illusion of Thinking,“whether current large reasoning models (LRMs) truly perform logical reasoning or if they merely simulate it through pattern recognition learned from training data



Recap



Machine Learning (ML)

Subset of AI -> focuses on algorithms -> learn and improve from data without being explicitly programmed for every scenario.

Key Concepts in Machine Learning:

- **Supervised Learning**:
  - Learn from labeled training data to make predictions on new, unseen data.
  - Map inputs to desired outputs based on example input-output pairs.

  Examples: Email spam detection, image classification, medical diagnosis
  - **Common algorithms**: Linear regression, decision trees, support vector machines, random forests

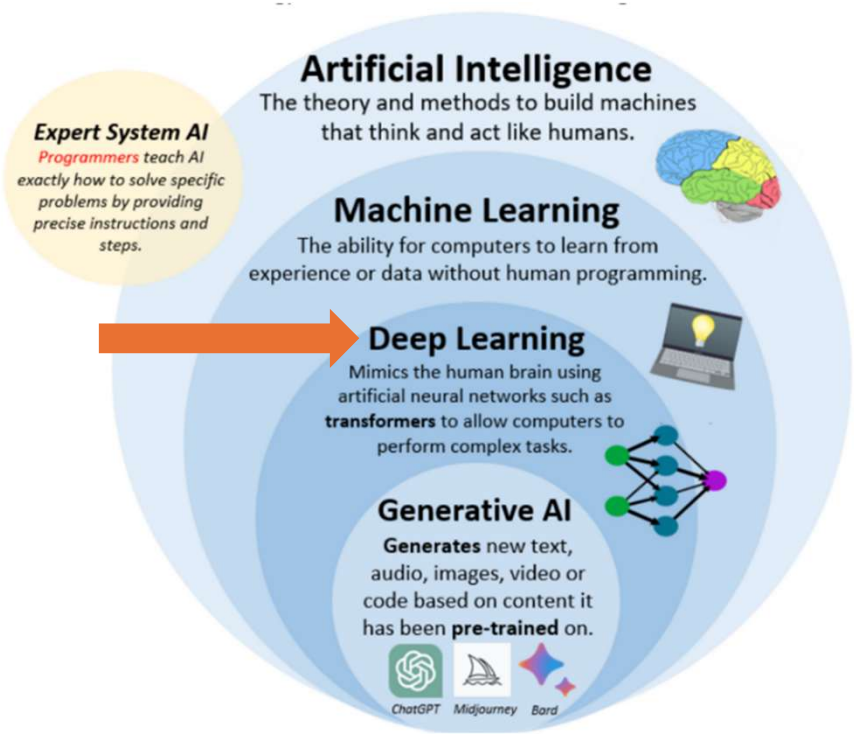
- **Unsupervised Learning**:
  - Find hidden patterns in data without labeled.
  - Discovers structure in data where the desired output is unknown.

  Examples: Customer segmentation, anomaly detection
  - **Common algorithms**: K-means clustering, hierarchical clustering, principal component analysis

- **Reinforcement Learning**:
  - learn through interaction with an environment,
  - Receiving rewards or penalties for actions taken.

  Examples: Game playing, robotics, autonomous vehicles, trading systems
  - **Common algorithms**: Q-learning, policy gradients, actor-critic methods

Recap



Deep Learning (DL)

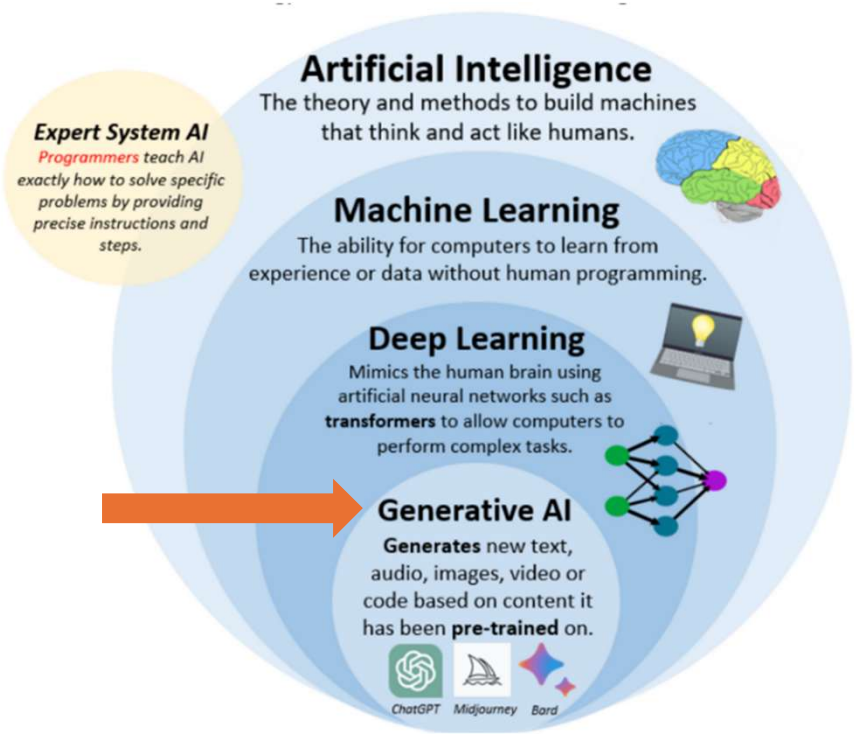
Subset of Machine Learning that uses artificial neural networks with multiple layers ("deep") and understand complex patterns in data.

- **Neural Network Fundamentals:**
  - Inspired by the structure of the human brain
  - Consisting of interconnected nodes (neurons) that process and transmit information.
  - Deep learning networks contain many layers of these neurons, for complex data and task.

Key Architectures in Deep Learning:

- **Feedforward Neural Networks**:
  - Information flows in one direction from input to output
  - Suitable for basic classification and regression tasks.
- **Convolutional Neural Networks (CNNs)**:
  - Specialized for processing grid-like data
  - Work fox ex : images, using convolutional layers that detect local features like edges and textures.
- **Recurrent Neural Networks (RNNs)**:
  - Designed for sequential data, with connections that create loops allowing information to persist,
  - For language processing and time series analysis.
- **Long Short-Term Memory (LSTM) Networks**:
  - Type of RNN that can learn long-term dependencies
- **Transformer Networks**:
  - Attention-based architectures
  - Revolutionized natural language processing, forming the basis for models like GPT and BERT.
- **Generative Adversarial Networks (GANs)**:
  - Consist of two competing networks that learn to generate realistic synthetic data, widely used for image generation and data augmentation.

Recap



Generative AI (GenAI)

Subset of Deep Learning that uses artificial Deep neural networks with multiple layers in generative manner (next word in the sequence).

- **Large Language model**
- **Foundation model**
- **Pretrained/Post-training/Re-Training**
- **Fine-tunning**
- **RAG**
- **Agentic architecture**

About ME

Trying to make this world a better place to live

AMBASSADOR

Gartner
Peer Community

COSCUP
開 源 人 年 會

AI Simplified

AWSSimplified

Preet Sharma

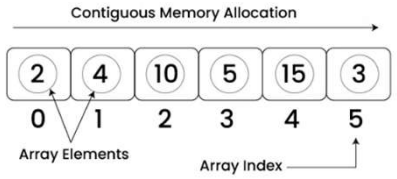
Problem Solver, Solutions Architect | Helping Organizations with Value-Innovation Solutions | C...

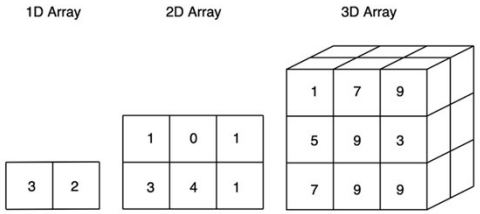
Lets start with the Basic

Why Mathematics

- Everything is Binary
- Liner algebra provides language for representing and manipulating data in high-dimensional spaces
- Probability and statistics offer frameworks for handling uncertainty and making inferences from data
- Calculus enables optimization of complex systems

Mathematics (liner algebra)





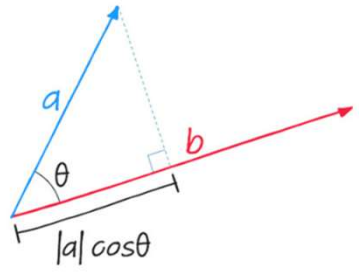

```
# 1D array
arr = [1, 2, 3]

# 2D array
arr_2d = [[1, 2, 3], [4, 5, 6]]

# 3D array
arr_3d = [[[1, 2], [3, 4]], [[5, 6], [7, 8]]]
```

Array

- Ordered collections of elements
- Arrays can have any number of dimensions
- Don't necessarily follow mathematical rules for operations.



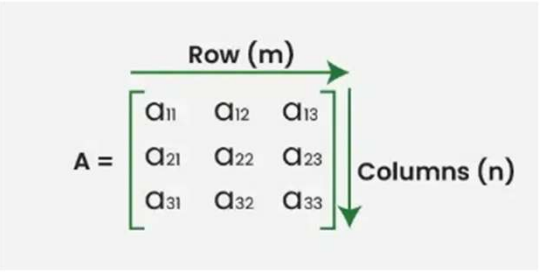
$$a \cdot b = |a|\,|b| \cos\theta$$

$$\vec{A} = 5\hat{\imath} + 8\hat{\jmath}$$

$$\vec{B} = 1\hat{\imath} - 3\hat{\jmath}$$

Vector

- One Dimensional Array
- Represents a mathematical vector
- Has both magnitude and direction in mathematical contexts
- Support mathematical operations like dot product, cross product, and magnitude calculations



$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$$

Matrix

- Two Dimensional Array in rows and column
- Specific math rules for operations
- Liner algebra ops (Multiply, determinant, inverse, eigenvalues)

Mathematics (liner algebra)

Vector operations

Addition $\vec{A} + \vec{B} = [A_1 + B_1, A_2 + B_2, A_3 + B_3, , , , , , ,]^T$

Scalar Multiplication $\vec{A}.B = [A_1.B, A_2.B, A_3.B, , , , , , ,]^T$

Dot product $\vec{A}.\vec{B} = [A_1B_1, A_2B_2, A_3B_3, , , , , , ,]^T = \left\|\vec{A}\right\| \cdot \left\|\vec{B}\right\| \cdot \cos(\theta)$ *θ is angle between vectors

$\left\|\vec{A}\right\|$ represent to **norm OR magnitude.** It is the length or size of the vector in geometry. The norm represents the distance from the origin (0,0,0…) to the point represented by the vector coordinates. It is always non negative scalar value.

There are three types of norms :

L1 norms (Manhattan norms) $\|\vec{A}\|_1 = |a_1| + |a_2| + \ldots + |a_n|$

Sum of absolute values of all components, ex. Vector: A= [3, -4, 5] = |3|+|-4|+|5| = 3+4+5 = 12

ex: Major the taxi cab block distance , you can move along street and not diagonal shortcut

L2 norms (Euclidean norms) $\|\vec{A}\| = \sqrt{(a_1^2 + a_2^2 + a_3^2 + \ldots + a_n^2)}$

Square root of the sum of squared components, ex. Vector: A= [3, -4, 5] $= \sqrt{(3^2 + (-4)^2 + 5^2)} = \sqrt{(9 + 16 + 25)} = \sqrt{50} \approx 7.071$

ex: Straight-line Euclidean distance - the shortest path "as the crow flies.

L∞ Norm (Maximum norm) $\|\vec{A}\|_\infty = \max(|a_1|, |a_2|, \ldots, |a_n|)$

Maximum absolute value among all components ex. Vector: A= [3, -4, 5] = max(|3|, |-4|, |5|) = max(3, 4, 5) = 5

completely determined by maximum values

Mathematics (liner algebra)

Example of usage:

Feature vectors

Numerical representation of an object or data point, where each dimension represents a specific measurable characteristic or "feature.“

House Feature Vector: [2100, 3, 2, 15, 1, 0]

Where each position represents:

•Position 0: Square footage (2,100 sq ft)
•Position 1: Number of bedrooms (3)
•Position 2: Number of bathrooms (2)
•Position 3: Age in years (15)
•Position 4: Has garage (1 = yes, 0 = no)
•Position 5: Has pool (0 = no, 1 = yes)

Email Feature Vector: [0.15, 847, 12, 3, 0, 1]

Representing:

•Position 0: Ratio of capital letters (15%)
•Position 1: Total word count (847)
•Position 2: Number of exclamation marks (12)
•Position 3: Number of dollar signs (3)
•Position 4: Contains "urgent" (0 = no)
•Position 5: Contains suspicious links (1 = yes)

Similarity computation

Used to identify similarity between two data points using cosine similarity

Movie Recommendation Example

- **Movie A (Action Thriller):** [4, 1, 5, 2, 3]
- **Movie B (Romantic Comedy):** [1, 5, 2, 1, 4]

Feature positions represent:

- Position 0: Action scenes (1-5 scale)
- Position 1: Romance level (1-5 scale)
- Position 2: Suspense level (1-5 scale)
- Position 3: Comedy level (1-5 scale)
- Position 4: Drama level (1-5 scale)

Cosine Similarity Calculation

Step 1: Dot Product

$$A \cdot B = (4\times1) + (1\times5) + (5\times2) + (2\times1) + (3\times4) = 4 + 5 + 10 + 2 + 12 = 33$$

Step 2: Magnitudes

$$|A| = \sqrt{(4^2 + 1^2 + 5^2 + 2^2 + 3^2)} = \sqrt{(16 + 1 + 25 + 4 + 9)} = \sqrt{55} \approx 7.42$$

$$|B| = \sqrt{(1^2 + 5^2 + 2^2 + 1^2 + 4^2)} = \sqrt{(1 + 25 + 4 + 1 + 16)} = \sqrt{47} \approx 6.86$$

Step 3: Cosine Similarity

Similarity = $33 / (7.42 \times 6.86) \approx 33 / 50.9 \approx$ **0.65**

Range: -1 to 1 (where 1 = identical, 0 = orthogonal, -1 = opposite)

Result:

0.65 indicates moderate similarity ,These movies share some common elements but are quite different genres

Dimensionality

Reduce dimension

Image Data :

- **Grayscale Image (28×28 pixels):**
  - Dimensionality: **784**
  - Each pixel is one dimension (0-255)
  - Feature vector: [$pixel_1$, $pixel_2$, ..., $pixel_{784}$]
- **Color Image (224×224×3):**
  - Dimensionality: **150,528**
  - Three color channels (RGB) per pixel
  - Extremely high-dimensional data

Customer Profile Text :

- Dimensionality: ~**50**
- Features: age, income, purchase_history, click_rates, etc.
- Relatively low-dimensional, dense data

The Curse of Dimensionality

As dimensions increase:

•**Low (1-10):** Easy to visualize and process
•**Medium (100-1000):** Manageable with good algo
•**High (10,000+):** Requires dimensionality reduction techniques
•**Very High (millions):** Computationally challenging, sparse data

Mathematics (Matrix)

Matrices are rectangular arrays of numbers that can represent systems of linear equations.
Matrix Notation: An m×n matrix A has m rows and n columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Matrix Addition:
Adding corresponding elements of matrices of the same dimensions:

A = [3 1] B = [1 5]
[2 4] [6 2]

C = A + B = [3+1 1+5] = [4 6]
[2+6 4+2] [8 6]

A = [1 -2 3] B= [2 1 -3]
[4 0 -1] [-1 3 4]
[2 5 6] [0 -2 1]

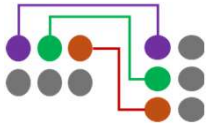
C=A+B = [1+2 -2+1 3+(-3)] = [3 -1 0]
? [4+(-1) 0+3 -1+4] [3 3 3]
[2+0 5+(-2) 6+1] [2 3 7]

Key Rule# Same Dimensions Required

Matrix Multiplication:
The fundamental operation for combining transformations:

A (2X3)= [1 2 3] B (3X2)= [7 8]
[4 5 6] [9 10]
[11 12]

C (2X2) =[58 64]
[139 154]



Step-by-Step Calculation
Element C_{11} (row 1, col 1):
(1×7) + (2×9) + (3×11) = 7 + 18 + 33 = 58
Element C_{12} (row 1, col 2):
(1×8) + (2×10) + (3×12) = 8 + 20 + 36 = 64
Element C_{21} (row 2, col 1):
(4×7) + (5×9) + (6×11) = 28 + 45 + 66 = 139
Element C_{22} (row 2, col 2):
(4×8) + (5×10) + (6×12) = 32 + 50 + 72 = 154

Number of columns in A must equal **Number of rows in B.**

Matrix Transpose:
Flipping a matrix along its diagonal:

A(2X3) = [1 2 3]
[4 5 6]

A^T = [1 4]
[2 5]
[3 6]

Rule: Rows become columns, columns become rows

Original Matrix (Student X Subject) #

	Math	Science	English
Alice	[85	92	78]
Bob	[76	88	85]
Carol	[91	79	92]

Transpose (Subject X Student) #

	Alice	Bob	Carol
Math	[85	76	91]
Science	[92	88	79]
English	[78	85	92]

Matrix Inverse:
Flipping a matrix along its diagonal:

A (2X2) = [3 1]
[2 4]

inverse: $A^{-1} = (1/\det(A)) \times [d - b] [-c a]$
Determinant
det(A) = (3×4) - (1×2) = 12 - 2 = 10

= (1/10) × [4 -1] [-c a]

= [0.4 -0.1] [-2 3]

= [-0.2 0.3]
[0.4 -0.1]

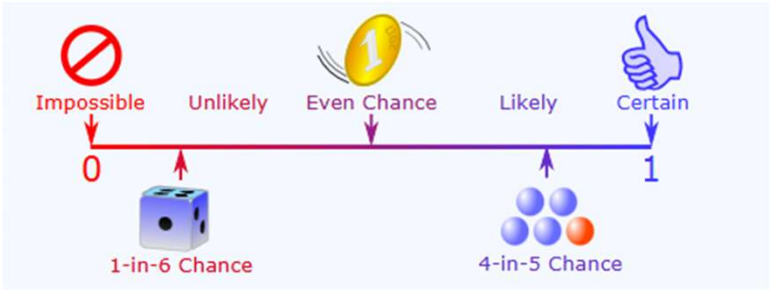
Inferential Statistics and Probability

Inferential statistics : The process of inferring insight from the sample data. Estimation can be made using the population data from the sample data, but not find the exact value. Only can make reasonable estimation with the limited level of certainty.

"Welcome to the world of uncertainty "

Basic definition of Probability: **Probability** is a mathematical measure of the **likelihood** that an event will occur , in given sample space.
It is expressed as a number between 0 and 1.

Probability Line



$$\text{Probability of an event happening} = \frac{\textit{Number of ways it can happen}}{\textbf{Total number of outcomes}}$$

Inferential Statistics and Probability

Terminology:

- Sample Space: all the possible outcomes of an experiment.
- Sample Point: just one of the possible outcomes
- Trial: A single performance of an experiment.
- Outcome: A possible result.
- Experiment: a repeatable procedure with a set of possible results.
- Event: one or more outcomes of an experiment

Example 1 : Roll a Dice what is the probability to get 6 ? =1/6
Example 2 : Toss a coin and what is the probability to get Head ? = ½
Example 3 : 5 marbles in a bag: 4 are blue, and 1 is red. What is the probability that a blue marble gets picked? =4/5

Ready for Hands On :

Q1 : How many times a "double – Both dice have same number" comes up when throwing 2 dice?

http://www.mathopolis.com/questions/q.php?id=700&site=1&ref=/data/probability.html&qs=700_701_702_1475_1476_1477_2175_2176_2177_2178

Inferential Statistics and Probability

Types of events “

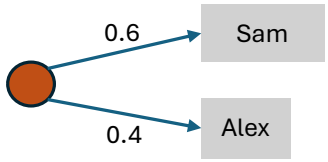
- Independent (each event is not affected by other events)
  - You toss a coin three times and it comes up "Heads" each time ... what is the chance that the next toss will also be a "Head"?

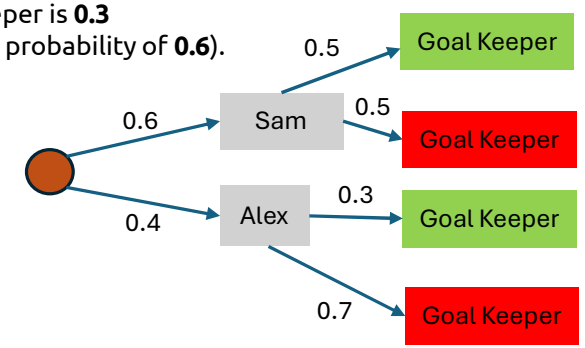
- Dependent (also called "Conditional", where an event is affected by other events)

  After taking one card from the deck there are less cards available, so the probabilities change!
  - For the 1st card the chance of drawing a King is 4 out of 52
  - But for the 2nd card :
    - If the 1st card was a King, then the 2nd card is **less** likely to be a King, as only 3 of the 51 cards left are Kings
    - If the 1st card was **not** a King, then the 2nd card is slightly **more** likely to be a King, as 4 of the 51 cards left are King

  You are off to soccer, and love being the Goalkeeper, but that depends who is the Coach today:
  - with Coach Sam your probability of being Goalkeeper is **0.5**
  - with Coach Alex your probability of being Goalkeeper is **0.3**

  Sam is Coach more often ... about 6 of every 10 games (a probability of **0.6**).





- Mutually Exclusive (events can't happen at the same time)
  - Turning left or right are Mutually Exclusive (you can't do both at the same time)
  - Heads and Tails are Mutually Exclusive
  - Kings and Aces are Mutually Exclusive

Inferential Statistics and Probability

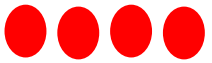
“House Always Wins”

Three Stage Probability Lifecycle

- Find all the possible Outcome
- Find likelihood (probability) of the combination
- Use probability to estimate Profit/Loss/Decision driver

Inferential Statistics and Probability

“Blue Ball /Red ball” : A bag is having 5 Balls (2 Blue and 3 Red). Each participant need to get a ball , note it color and put it back in the bag. Every participant do it 4 times. Participant who get red ball all four times will receive 150 INR and for any other result participant needs to pay back 10 INR. Would you like to play ???

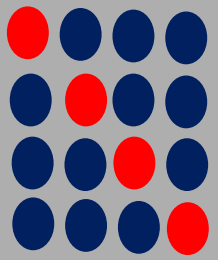
 = 150 INR

Any Other combination = -10 INR

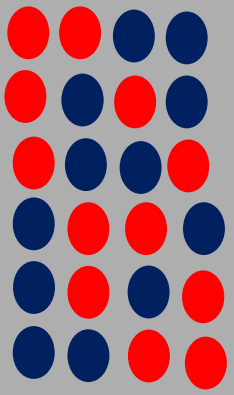
Step 1 # Find all the possible outcome

Possible combination
4 Blue Balls
0 Red balls

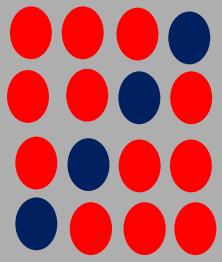
Possible combination
3 Blue Balls
1 Red balls



Possible combination
2 Blue Balls
2 Red balls



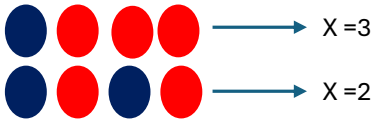
Possible combination
1 Blue Balls
3 Red balls



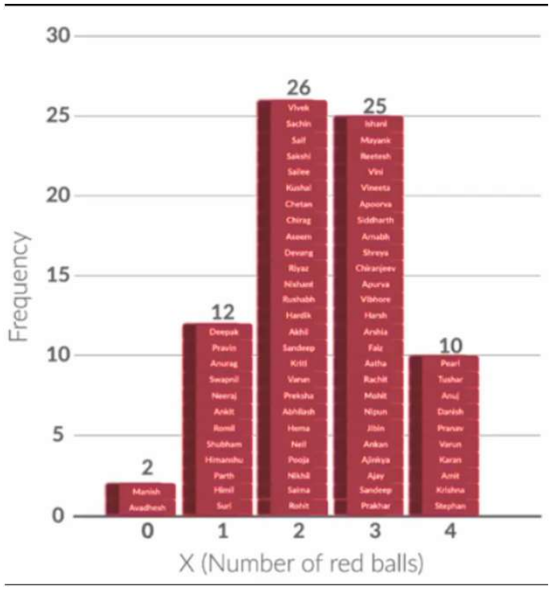
Possible combination
0 Blue Balls
4 Red balls

Inferential Statistics and Probability

Step 2 : Estimate likelihood for relative combination
Quantify Outcome X = Number of Red balls



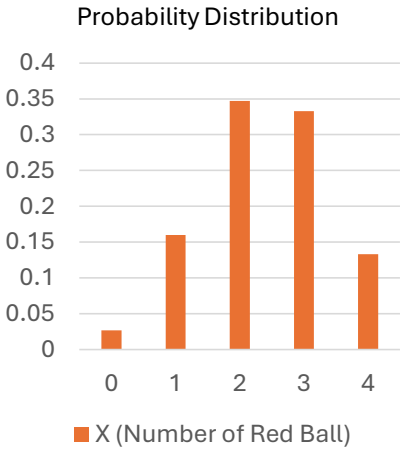
X is random number in statistical language
It is an outcome between outcome of an experiment and a natural number



- Experiment for 75 people.
- what's the probability that people get 2 Red balls (X=2) ?

Probability distribution for each likelihood.

X	P(X)
0	0.027
1	0.160
2	0.347
3	0.333
4	0.133



Inferential Statistics and Probability

So lets see in long run this game would be profitable to the house or to the player.
Lets say 1000 people are playing the came.

X	P(X)	P(1000)
0	0.027	1000X0.027 =27
1	0.160	1000X0.16=160
2	0.347	1000X0.347=347
3	0.333	1000X0.333=333
4	0.133	1000X0.133=133

How many time there is red ball ?
Expected no of times red ball = 0X27+1X160+2X347+3X333+4X133= 2385 times
Average number of red ball for 1 game =2385/1000 = 2.385

Expected Value EV=x_1*P(X= x_1)+x_2*P(X= x_2)+x_3*P(X= x_3)+x_4*P(X= x_4)+..... x_n*P(X= x_n)

EV=0*P(X=0) + 1*P(X=1)+ 2*P(X=2)+ 3*P(X=3)+ 4*P(X=4)
i.e. EV =0*(0.027)+1*(0.16)+2*(0.347)+3*(0.333)+4*(0.133) =2.385

So Any Idea now House is in Profile or Loss ???

X can take two value +150 and -10
P(X=+150) i.e. P(4 red balls) = 0.133
P(X=-10) i.e. P(0,1,2,3 red balls)=0.027+0.160+0.347+0.333=0.867

So EV = (150 X 0.133) + (-10 X 0.867) =+11.28
That means on average person will win **11.28 INR** from the Game.

So will House loose or will make money ?

For House to Win EV should be in negative , for that below are the possible options :

- Decrease reward money from 150 to 100
- Increase penalty from 10 to 50

Will it bring EV as negative ?

$$EV(X) = \sum_{i=1}^{i=n} x_i * P(X = x_i)$$

Inferential Statistics and Probability

Rules of Probability
Addition rule of probability

Addition Rule 1: When two events, A and B, are mutually exclusive, the probability that A or B will occur is the sum of the probability of each event.

P(A or B) = P(A) + P(B)

Ex: A single 6-sided die is rolled. What is the probability of rolling a 2 or a 5? (Mutely exclusive events)

Additional Rule 2: When two events, A and B, are non-mutually exclusive, the probability that A or B will occur is:

P(A or B) = P(A) + P(B) – P(A and B)

Ex1: In a math class of 30 students, 17 are boys and 13 are girls. On a unit test, 4 boys and 5 girls made an A grade. If a student is chosen at random from the class, what is the probability of choosing a girl or an A student?

Probabilities: P(girl or A) = P(girl) + P(A) – P(girl and A)

$$= \frac{13}{30} + \frac{9}{30} - \frac{5}{30} = \frac{17}{30}$$

Ex2: On New Year's Eve, the probability of a person having a car accident is 0.09. The probability of a person driving while intoxicated is 0.32 and probability of a person having a car accident while intoxicated is 0.15. What is the probability of a person driving while intoxicated or having a car accident?

Probabilities : P(Intoxicated or accident) = P(intoxicated) +P(accident) - P(intoxicated and accident)

= 0.32 + 0.09 - 0.15

= 0.26

Inferential Statistics and Probability

Rules of Probability
Multiplication rule of probability (For compound events)

Multiplication Rule 1: When two events, A and B, are independent, the probability of both occurring is:
P(A and B) = P(A) · P(B)

Ex : A drawer contains one pair of socks with each of the following colors: blue, brown, red, white and black. Each pair is folded together in a matching set. You reach into the sock drawer and choose a pair of socks without looking. You **replace** this pair and then choose another pair of socks. What is the probability that you will choose the **red pair of socks both times**?
P(red) =1/5
P(red and red) = p(red) .p(red) = 1/5.1/5 =1/25

Ex : A coin is tossed and a single 6-sided die is rolled. Find the probability of landing on the head side of the coin and rolling a 3 on the die.
P(head) =1/2
P(die side =3)=1/6
P(head and 3) = P(head).P(3) = 1/2.1/6 = 1/12

Inferential Statistics and Probability

Probability without experiments

Binominal distribution

(probability distribution that models the number of successes in a fixed number of independent trials, where each trial has the same probability of success.)

Key characteristics

- **Fixed number of trials (n)**
- **Only two outcomes** per trial (success/failure, red/blue)
- **Constant probability** of success (P) for each trials
- **Independent trials** (one doesn't affect another's)

Formula : $P(X = k) = C(n,k) \times p^k \times (1-p)^{(n-k)}$

Where

- X = Total number of successes
- k = specific number of successes we want
- n = total number of trials
- p = probability of success on each trial
- C(n,k) = combinations = n!/(k!(n-k)!)

Binomial Distribution Applicable	Binomial Distribution Not Applicable
Tossing a coin 20 times to see how many tails occur	T ossing a coin until a heads occurs
Asking 200 randomly selected people if they are older than 21 or not	Asking 200 randomly selected people how old they are
Drawing 4 red balls from a bag, putting each ball back after drawing it	Drawing 4 red balls from a bag, not putting each ball back after drawing it

Ex1 : Flip a fair coin **10 times**. What's the probability of getting **exactly 6 heads**?

Here n =10 trials , k= 6 success (head), p=0.5 (probability of head)

$P(X = 6) = C(10,6) \times (0.5)^6 \times (0.5)^4$

$P(X = 6) = 210 \times 0.015625 \times 0.0625$

$P(X = 6) = 0.205$

$P(X = 6) = 20.5\%$

Ex2 : A factory produces items with **95% success rate**. In a **batch of 20 items**, what's the probability that **exactly 18** are good?

Here n =120 trials , k= 18 success (good), p=0.95 (probability)

$P(X = 18) = C(20,18) \times (0.95)^{18} \times (0.05)^2$

$P(X = 18) = 190 \times 0.397 \times 0.0025$

$P(X = 18) = 0.189$

$P(X = 18) = 18.9\%$

Concept of Hypothesis

Hypothesis is a testable statement or educated guess about a relationship between variables or a characteristic of a population that can be evaluated through statistical analysis.

- Null Hypothesis (H_0)
    - Definition : Default assumption or “status quo”
    - Purpose : state that there is no effect , no difference or no relationship
    - Example : “New drug has no effect on blood pressure”

- Alternative Hypothesis (H_1 or H_a)
    - Definition: What we want to prove or investigate
    - Purpose: States there is an effect, difference, or relationship
    - Example: "The new drug reduces blood pressure“

Example 1: Drug Effectiveness
Research Question: Does the new blood pressure medication reduce systolic blood pressure?
- **H_0:** μ = 140 mmHg (no change in blood pressure) -- μ is population data
- **H_1:** μ < 140 mmHg (medication reduces blood pressure)

Study Design**:** Give medication to 100 patients, measure blood pressure after 4 weeks

Example 2: Customer Satisfaction
Research Question: Does the new customer service training improve satisfaction scores?
- **H_0:** μ_after = μ_before (training has no effect)
- **H_1:** μ_after > μ_before (training improves satisfaction)

Study design : give training to 100 Employee related to customer satisfaction training.

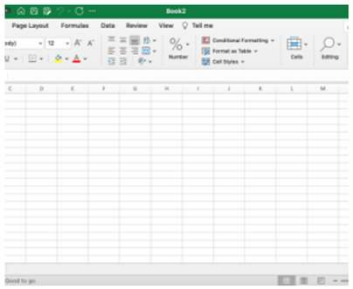
Research Question: Do students learn better with interactive vs. traditional lectures?
- **H_0:** μ_interactive = μ_traditional (no difference in test scores)
- **H_1:** μ_interactive > μ_traditional (interactive method produces higher scores)

Study Design: Randomly assign 200 students to each teaching method

DATA

Structured Data

- information organized in a predefined, consistent format, making it easy to store, retrieve, and analyze
- Relational Databases e.g. MySQL, PostgreSQL Columnar for retrieval optimized e.g. Google Bigtable, Cassandra In-memory for faster access e.g. Redis, Memcached



01 | Spreadsheets, CSV

02 | Enterprise Data Warehouses

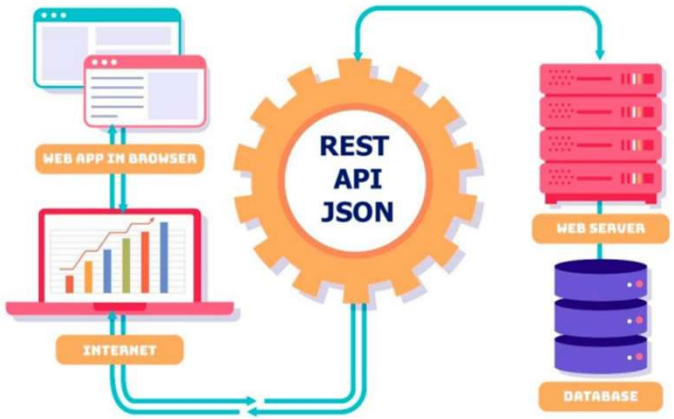
03 | ERP e.g. SAP, Oracle, Zoho, Tally

04 | CRM e.g. Salesforce, Zendesk, Adobe

05 | HR mgmt. eg. Workday, Ceridian, Oracle

Semi Structured Data

- Type of data that falls between structured and unstructured data. It possesses some organizational properties, like tags or markers, but doesn't conform to the rigid structure of a relational database
- “Document” Databases e.g. MongoDB, CouchDB
  Graph Databases e.g. Ne04j, Amazon Neptune Wide-Column Databases e.g. Cassandra, HBase



01 | Data exchange between server and client in apps, webservices

02 | Configuration files in software applications

03 | Application log files w/ info about events, errors, user interactions

04 | Used for encoding geographical data e.g. GeoJSON

05 | IoT devices exchange data between sensors, devices, and servers.

Text

- A local file system or Network-attached storage (NAS)
  Object storage e.g. Amazon S3, Google Cloud Storage
  Content Management Systems e.g. Wordpress, Drupal

01 | Webpages and PDFs

02 | Books, Novels & Literature

03 | Social Media Platforms e.g. Twitter, Facebook,

04 | Chat and Forums e.g. Whatsapp, Stackoverflow, Quora

05 | Research Publications, arXiv.

Images

- A local file system or Network-attached storage (NAS)

  Object storage e.g. Amazon S3, Google Cloud Storage

01 | User Generated Content from Social Media platforms

02 | Satellite imagery e.g. from NASA or commercial providers

03 | Medical Imaging e.g. CT-Scans, MRIs, X-Rays

04 | Frames extracted from videos

Video

- A local file system or Network-attached storage (NAS)
  Cloud Video storage e.g. AWS Elemental MediaStore
  Media Asset Management Systems e.g. Kaltura, Brightcove

01 | Video streaming platforms like YouTube

02 | OTT platforms like Netflix, Amazon Prime

03 | EdTech platforms e.g. Khan academy

04 | Broadcast media from TV, Cable, Satellite

05 | Movies, Documentary and other entertainment

Audio

- A local file system or Network-attached storage (NAS)
  Cloud storage e.g. AWS S3, Google Cloud Storage,
  Azure Blob Audio Databases for efficient retrieval e.g.
  AcoustID, MusicBrainz

01 | Streaming platforms like Spotify, Wynk

02 | Podcasts & Audiobooks

03 | Webinars and Event livestreams

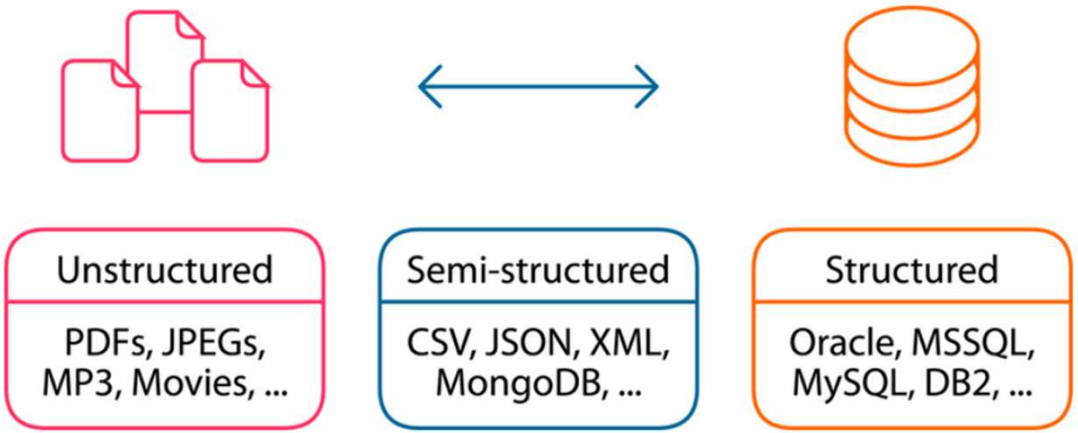
04 | Field Recordings of nature sounds, animals, machinery, cities

05 | Voice assistants like Amazon Alexa, Google Assistant

Data Formats

- A local file system or Network-attached storage (NAS)
  Cloud storage e.g. AWS S3, Google Cloud Storage,
  Azure Blob Audio Databases for efficient retrieval e.g.
  AcoustID, MusicBrainz

- In Data Science, ML algorithms are typically applied to structured data.
- In AI, ML algorithms are typically applied to text, video, image, audio
- Real world deployments often rely on a pipeline of data pre-processing, data transformations, ML algorithms, post-processing



Data Formats

Unstructured Data

The university has 5600 students. Shaun (ID Number: 160801), 18 years old Communication study. Linh with ID number 160802, majoring in Accounting and is 20 years old. Ahmed from Psychology study program, 19 years old, ID number 160803.

Semi-Structured Data

```
<University>
 <ID Number="160801">
  <Name="Shaun">
  <Age="18">
  <Program="Communication">
 <ID Number="160802">
  <Name="Linh">
  <Age="20">
  <Program="Accounting">
..... </University>
```

Structured Data

ID	Name	Age	Program
160801	Shaun	18	Communication
160802	Linh	20	Accounting
160803	Ahmed	19	Psychology